



City Research Online

City, University of London Institutional Repository

Citation: Crook, J., Bellotti, T., Mues, C. and Fuertes, A-M. ORCID: 0000-0001-6468-9845 (2019). Preface to the papers on 'Credit risk modelling'. Journal of the Royal Statistical Society Series A, 182(4), pp. 1139-1142. doi: 10.1111/rssa.12525

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/25199/>

Link to published version: <http://dx.doi.org/10.1111/rssa.12525>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Preface

Credit risk modelling involves the estimation and prediction, usually, of the probability that an applicant for credit or an account holder will not repay as agreed within a time period. It is one of the most successful and widespread applications of statistical methodologies in finance. Almost every adult who has a bank account or a loan has had a risk of default score attributed to them at some time using a credit risk model and in most cases the scores have been updated monthly. Academic and practitioner research in the area has increased rapidly following changes in legislation concerning, for example, provisions that lenders must hold against default (under IFRS9 or CECL in the US) or the amount of capital that banks are required to hold to protect depositors in the event of severe but reasonable deteriorations in an economy (Basel II/III), or changes in technology (facilitating the entry of fintechs) or the availability of transactions data. This increase in attention was manifested at the Credit Scoring and Credit Control conference XV organised by the Credit Research Centre at the University of Edinburgh in August 2017 at which many of the papers in this Collection were presented.

This Collection contains eight papers that broadly fall into four areas: improved statistical methodologies to enhance or assess predictive accuracy; the incorporation of network effects as predictors that enhance predictive accuracy; the parameterisation of exposure at default and loss given default distributions and the implications of prohibiting the use of a minority characteristic, in this case gender, from scoring models.

An issue concerning credit risk modellers is the highly imbalanced nature of the population and development samples: ~~there are~~ typically less than 4% of a population of account holders ~~who~~ do not pay as agreed and this can be under 1% in the case of mortgages (Thomas et al 2017). There are two potential problems. The maximum likelihood estimator of the parameters of a logistic regression is biased in small samples and the degree of bias depends on the number of instances in the less frequently populated class. The second problem is that the rate at which the probability of class membership approaches 1 may differ from the rate at which it approaches 0 in which case a symmetric link function may not be appropriate. The paper by Ogundimu compares alternative methods that have been proposed in the literature to gain more accurate predictions of the probability of default using a sample of credit card accounts from Taiwan. He finds that, for his dataset, if AUROC is the measure of discrimination then a SMOTE (Synthetic Minority Oversampling Technique) should be preferred over the other methods considered and that penalized logistic regression and Firth's method (Firth 1993) with additional covariates were superior to logit, the Firth method and to Generalized extreme value regression (Calabrese and Osmetti 2013). If AUPRC is the accuracy measure then ridge regression and random oversampling examples are preferred.

Estimating the probability of default (PD) in a bivariate context has many applications. Calabrese, Osmetti and Zanin propose a bivariate probability of default model where the PD is estimated conditional on whether or not the applicant has defaulted on a loan as recorded by a credit bureau. They deduce the joint probability of default in terms of a copula function when the generalized extreme value function (GEV) is used to model the marginal default probabilities. They specify an estimation procedure for the GEV model for a chosen copula. Using simulated data to study the properties of the model they find that the proposed model gives more accurate predictions than alternative single equation models or a bivariate probit model. An illustration of the model is provided with a Peer-2-Peer loan level dataset.

Assessing the predictive accuracy of classifiers has been a research topic for many years (Thomas et al 2017). Coolen-Maturi and Coolen consider the use of non-parametric predictive inference (NPI) for this purpose. NPI is a statistical method based on Hill's assumption (Hill 1968) concerning the prediction as to which interval an as yet unobserved value belongs, conditional on past observations but without making any assumptions about the distribution of the observations. This assumption does not allow precise probabilities to be derived but does allow upper and lower bounds on the distribution of probabilities to be derived. Coolen-Maturi and Coolen argue that ~~that~~ the distributions of existing borrowers is, in practice, unknown and apply these concepts to the situation where a scoring model yields a predicted ordered class from more than three such classes which indicate decreasing degrees of credit worthiness. They show the conditions that the points on the NPI lower Receiver Operating Curve (ROC) and upper ROC hypersurfaces have to satisfy and also the envelopes of these bounds. They do likewise for the hypervolumes under the ROC hypersurfaces. They illustrate the use of these bounds for evaluating the performance of classifiers using two datasets.

The next two papers incorporate different aspects of network information into PD modelling. Information on transactions between SMEs may be expected to increase one's ability to predict the probability that a company would default on its loans. If a supplier or a customer has difficulties then the firm of interest may also be in trouble. But research on this topic has been hampered by lack of data. Tosetti, Moscone and Lycett use information relating to a large number of inter-firm transactions to investigate whether this type of information may enhance predictive accuracy. They consider network characteristics such as the number of companies from which transactions are received, the number to which transactions are sent, the amount of inward transactions and first order neighbourhood characteristics – characteristics of companies from whom income is received. They find that the inclusion of total number and volume of inward and outward transactions significantly increases predictive accuracy of PD models. In the second paper in this group, Tobback and Martens consider direct networks, where transactions occur directly between the network members, and implied networks where members are those who transferred or received money from the same entity. They show a method to identify network membership. They find that traditional models outperformed the network models and are complimentary to them. They also find that being part of a direct network increases accuracy more than being part of an indirect network: being connected to a defaulter increases the probability of default. At high scores, models using payment data in a direct network have more defaulters than at the highest scores of a traditional model. Having more months' worth of transactions data increases predictive accuracy.

The third group of papers address issues relating to capital modelling, both regulatory capital under the Basel Accords and economic capital. At the level of an account, loss given default (LGD) is the proportion of the exposure at the time of default that the lender never receives in the event of default (Bellotti and Crook 2012). LGD distributions have been observed to have a wide variety of shapes, but are usually concentrated at 0, 1 and are sometimes multimodal in between. This makes modelling and prediction of LGD very challenging and a large number of approaches have been attempted in the literature (Loterman et al 2012). Tomarchio and Punzo use a zero-one inflated mixture model with a three level multinomial model to classify defaulted accounts into three sets: (0), (1) and {0,1} and a finite mixture of distributions is used to model values in the set (0,1). To test hypotheses concerning the appropriate distributions they use two data sets; one from a European bank and one from the Bank of Italy. Unlike previous studies they allow multiple mixture distributions in (0,1). They find that almost all of their models had at least three mixture components and that no single model gave more accurate predictions in all cases.

Thackham and Ma are concerned with modelling the exposure at default (EAD) on which there has been much less empirical research than on LGD. Possibly influenced by practitioners much of the literature that does exist considers methods based on transformations of EAD (Leow and Crook 2016). These are the credit conversion factor, exposure at default factor and loan equivalent factor with the credit conversion factor being arguably the most popular. Thackham and Ma build a descriptive model of EAD without using the CCF transformation. They use a mixture model that includes the probability that the limit has fallen by the time of default and the value of the balance. The latter is assumed to follow a two component normal mixture model. Each is conditional on observed covariates. They find that their model has a good degree of predictive accuracy and provide insights into the appropriate statistical drivers of EAD.

The final paper in the issue by XXX (authors; add to references) tackles an issue of rapidly increasing academic activity: the occurrence of unintended segment bias in empirical credit risk models. The segment they consider is females. They investigate the correlation between default occurrence and gender using a Markov blanket that includes these two variables and others that turn out to be correlated with gender. They explore this further using correspondence analysis. They conclude that whilst gender is statistically significant in a scoring model its removal does not reduce predictive accuracy. But of course removing such a variable may alter the characteristics of individuals who are accepted for credit. Including gender in a model is found to increase the predicted probability of a female being classified as a low-risk borrower compared with a model that omits gender, whereas a model that omits gender yields a lower probability for a female to be accepted than a model that includes gender. However, in general the rejection rates for females are lower than for men so the removal of gender does not lead to the same outcome between the genders. They suggest a simple solution. ← AMF: do we need this last sentence? It does not say much.

We would like to thank all of the authors who submitted their papers to the Collection and for making the changes requested by the Guest Editors. We would also like to thank all of the referees who kindly gave up their time to review the papers submitted and we would like to thank the Joint Editor who oversaw our work as Guest Associate Editors.

References

- Bellotti, T and Crook, J (2012) Loss given default models incorporating macroeconomic variables for credit cards. *International Journal of Forecasting*, 28, 171-182.
- Calabrese, R. and Osmetti, S.A. (2013) Modelling small and medium enterprise loan defaults as rare events: the generalized extreme value regression model. *Applied Statistics*, 40(6), 1172-1188.
- Firth, D (1993) Bias reduction in maximum likelihood estimators. *Biometrika*, 80(1), 267-238.
- Hill, B.M. (1968) Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63, 677-691.
- Leow, M. and Crook, J. (2016) A new mixture model of the estimation of credit card exposure at default. *European Journal of Operational Research*, 236, 685-694.
- Loterman, G, Brown, I., Martens, D., Mues, C. and Baesens, B (2012) Benchmarking regression algorithms for loss given default modelling. *International Journal of Forecasting*, 28, 161-170.
- Thomas, L. Crook, J. and Edelman, D (2017) *Credit Scoring and its applications*. Philadelphia: SIAM.

Jonathan Crook
University of Edinburgh

Tony Bellotti
Imperial College, London

Christophe Mues
University of Southampton

Ana-Maria Fuertes
City University of London